Roll No.											

PRESIDENCY UNIVERSITY

BENGALURU

Mid - Term Examinations – March 2026

Date: 12- 03- 2026 **Time:** 09:30am – 11.00am

School: SOCSE	**Program:** B.Tech	
Course Code: CSD1712	**Course Name:** Statistical Foundations of Data Science	
Semester: IV	**Max Marks**: 50	**Weightage**: 25%

CO - Levels	**CO1**	**CO2**	**CO3**	**CO4**	**CO5**
Marks	**24**	**26**	**0**	**0**	**0**

Instructions:

(i) Read all questions carefully and answer accordingly.

(ii) Do not write anything on the question paper other than roll number.

Part A

Answer ALL the Questions. Each question carries 2marks. **5Q x 2M=10M**

1	Recall the role of Big Data in Health Sciences.	**2 Marks**	**L1**	**CO1**
2	Define spurious correlation.	**2 Marks**	**L1**	**CO1**
3	Define Polynomial Regression.	**2 Marks**	**L1**	**CO2**
4	What is the Bias–Variance Tradeoff?	**2 Marks**	**L1**	**CO2**
5	What is an Exponential Family of Distributions?	**2 Marks**	**L1**	**CO2**

Part B

Answer the Questions. **Total Marks 40M**

6.	**a.**	Demonstrate the concept of Big Data and its key characteristics.	**10 Marks**	**L2**	**CO1**
	b.	Explain the aim of high-dimensional statistical learning.	**10 Marks**	**L2**	**CO1**
Or					
7.	**a.**	Outline the rise of Big Data and dimensionality across different scientific domains.	**10 Marks**	**L2**	**CO1**
	b.	Illustrate noise accumulation and its effects in high-dimensional data.	**10 Marks**	**L2**	**CO1**

8.	**a.**	Experiment with Multiple Linear Regression and the Gauss–Markov Theorem.	**10 Marks**	**L3**	**CO2**
	b.	Make use of Residual Variance and Refitted Cross-validation.	**10 Marks**	**L3**	**CO2**
Or					
9.	**a.**	Utilize the concept of Iteratively reweighed least squares.	**10 Marks**	**L3**	**CO2**
	b.	Identify Ridge Regression, Penalized Least Squares, and Bayesian Interpretation.	**10 Marks**	**L3**	**CO2**